

ANALISIS PENGGUNA WHATSAPP TERHADAP KESALAHAN MENGIRIM PESAN TEKS MENGGUNAKAN METODE KLASIFIKASI

Finki Dona Marleny¹⁾, Mambang²⁾

¹⁾Sistem Informasi Universitas Sari Mulia

²⁾Teknologi Informasi Universitas Sari Mulia

Jl. Pramuka No.2 Banjarmasin

e-mail: finkidona@unism.ac.id¹⁾, mambang@unism.ac.id²⁾

ABSTRAK

Aplikasi perpesanan sederhana lintas platform yang memungkinkan untuk bertukar pesan, mulai dari pesan teks, gambar, video, serta panggilan video dan panggilan cepat tanpa biaya banyak digunakan oleh pengguna di Indonesia, salah satunya adalah Aplikasi Whatsapp. Platform tersebut banyak digunakan dan diminati semua kalangan, dalam percakapan pribadi maupun percakapan grup. Percakapan yang didominasi dengan menggunakan pesan teks sangat memungkinkan terjadinya kesalahan pengetikan dalam pengiriman pesan yang mengakibatkan kesalahan persepsi pembaca pesan tersebut. Berdasarkan kesalahan pengiriman pesan dan kesalahan dalam pengetikan pesan teks pengguna platform tersebut maka akan dilakukan analisa pengguna aplikasi Whatsapp di lingkungan Universitas XYZ menggunakan metode klasifikasi dengan algoritma Naïve Bayes dan KNN terhadap kesalahan mengirim pesan teks. Dari data yang didapat yaitu hasil klasifikasi menggunakan algoritma Naïve Bayes mendapatkan akurasi sebesar 75%, sedangkan klasifikasi menggunakan algoritma KNN mendapatkan akurasi sebesar 66,7%. Dari hasil klasifikasi tersebut performa algoritma Naïve Bayes lebih besar dari pada performa algoritma KNN.

Kata Kunci: Whatsapp, pesan, klasifikasi, KNN, Naïve Bayes

ABSTRACT

A simple cross-platform messaging application that allows you to exchange messages, ranging from text, picture, video, and speed-dialing messages without cost, is widely used by Users in Indonesia, one of which is the Whatsapp Application. The platform is widely used and sought after in all walks of life, in private conversations and group conversations. Discussions that are dominated by using text messages are very likely typographical errors in sending messages that use the reader's misperception of the message. Based on the mistake of sending messages and errors in typing text messages on this user platform, this study will analyze the Whatsapp application in XYZ University environment using classification methods with the Naïve Bayes algorithm and KNN on text message sending messages. From the data that can be studied, the results of classification using the Naïve Bayes algorithm get an accuracy of 75%, while the classification using the KNN algorithm gets an accuracy of 66.7%. From the results of this classification, the performance of the Naïve Bayes algorithm is higher than the performance of the KNN algorithm.

Keywords: Whatsapp, messaging, classification, KNN, Naïve Bayes.

I. PENDAHULUAN

Kebutuhan akan perpesanan sederhana lintas platform yang memungkinkan untuk bertukar pesan, mulai dari pesan teks, gambar, video, serta panggilan video dan panggilan cepat tanpa biaya merupakan kebutuhan yang tidak dapat dipisahkan bagi masyarakat era revolusi industri 4.0 saat ini, dimana biaya yang dibebankan tidak berupa biaya pulsa seperti pada layanan *Short Message Service* (SMS), dimana penggunaan aplikasi pesan lintas platform atau sering disebut aplikasi perpesanan menggunakan layanan data untuk berkirim pesan yang menjadi tren baru seiring era *smartphone*.

Banyak jenis dari Aplikasi *chat* yang ditawarkan dan memiliki fitur-fitur yang beragam, salah satu yang paling populer digunakan di Indonesia adalah Whatsapp [1], dengan kemudahan yang ditawarkan banyak pengguna aplikasi chatting di Indonesia memilih menggunakan Whatsapp baik untuk percakapan pribadi

maupun untuk percakapan grup, dalam hal ini aplikasi tersebut memiliki keunggulan dan kelemahan tersendiri, contohnya pada analisis beberapa responden sebenarnya tidak suka berada di grup whatsapp tertentu, namun karena alasan yang beragam tidak dapat meninggalkan percakapan grup tersebut, selain tentang privasi dan keamanan.

Beberapa kelemahan dalam aplikasi whatsapp adalah dalam kesalahan pengiriman pesan teks dan kesalahan pengetikan pada saat mengetik serta mengirimkan pesan, karena pada aplikasi whatsapp tidak memiliki fitur untuk memperbarui pesan yang salah dalam pengetikannya, sehingga pesan yang salah dalam pengetikan hurufnya hanya dapat di hapus keseluruhan pesan yang telah dikirimkan hal ini mendapatkan respon yang beragam dari responden.

Analisis sentimen yang juga sering disebut sebagai *opinion mining* adalah rangkaian proses dalam memberikan informasi dan hasil dari olahan suatu sistem yang difungsikan untuk mengklasifikasi antara kategori positif dan negatif pada keseluruhan sentimen yang dinyatakan dalam tiap dokumen [2]. Pemikiran umum text mining merupakan untuk mengetahui bagaimana cakupan atau topik dari permasalahan dalam teks [3]. Secara esensial Text mining sangat penting dalam melakukan analisis sentimen sebagai identifikasi terhadap emosional suatu pernyataan, sehingga banyak studi yang mengkaji dan menganalisis sentiment yang dilakukan sesuai dengan kategori yang akan diteliti [4].

Algoritma Naïve Bayes yang menggunakan teknik pembelajaran mesin berbasis probabilitas memiliki hasil pengukuran dengan nilai yang mendekati akurat dapat lebih baik dalam mengklasifikasi serta performansi yang tinggi baik digunakan dalam pengklasifikasian [5] k-Nearest Neighbor (kNN) termasuk kelompok *instance-based learning* yang dapat digunakan untuk proses klasifikasi. Algoritma ini mudah membaca data kelompok menjadi salah satu teknik *lazy learning*. kNN diperoleh dengan mencari kelompok *k* objek di dalam data *training* yang paling mirip kedekatannya dengan objek pada data testing [6].

Selanjutnya akan dilakukan analisa terhadap pengguna whatsapp di lingkungan Universitas XYZ menggunakan metode klasifikasi dengan algoritma Naïve Bayes dan KNN terhadap kesalahan mengirim pesan teks.

II. METODE PENELITIAN

Data yang digunakan dalam analisis pengguna Whatsapp ini berupa opini tentang pengguna aplikasi Whatsapp dalam kesalahan mengirim pesan pada pesan pribadi dan percakapan grup. Dari data yang tersedia, diambil sebanyak 150 peserta grup Whatsapp yang aktif di dalam percakapan grup tersebut. Opini yang digunakan yaitu 200 opini yang di klasifikasikan berupa 100 opini positif dan 100 opini negatif, data percakapan yang digunakan untuk proses klasifikasi adalah data percakapan grup selama satu tahun. Data yang digunakan berupa data untuk pembelajaran mesin dan data uji untuk evaluasi kinerja pengklasifikasian.

Berikut tahapan untuk analisis pengguna Whatsapp adalah:

1. Tahap *pre-processing* data

Pada tahapan ini, data percakapan grup selama satu tahun dikumpulkan untuk dapat dimasukkan ke dalam aturan dalam klasifikasi, tahapan awal data mentah dilakukan proses tokenizer, stemming, serta *stopwords*. Hasil dari tahapan ini menghasilkan fitur yang digunakan sebagai data pembelajaran mesin oleh Naïve Bayes dan kNN.

2. Tahap seleksi dengan fitur *Chi Square*

Selanjutnya tahapan seleksi dapat menggunakan seleksi fitur *Chi Square*. Langkah awal yang dilakukan adalah dengan menentukan tabel kontingensi pada masing-masing fitur dengan menghitung nilai seleksi fitur *Chi Square* dengan persamaan berikut [7]):

$$X^2(D, t, c) = \frac{(N_{00} + N_{11} + N_{10} + N_{01}) \times (N_{00}N_{11} - N_{10}N_{01})^2}{(N_{11} + N_{01}) \times (N_{11} + N_{10}) \times (N_{10} + N_{00}) \times (N_{01} + N_{00})} \quad (1)$$

pengamatan kebersesuaian (*goodness of fit*) menggunakan seleksi fitur Chi Square dari setiap kategori. Uji Chi Square dalam statistika di aplikasikan untuk menguji independensi dari 2 (dua) peristiwa. Sedangkan pada seleksi fitur berdasarkan teori statistika, dua peristiwa tersebut di antaranya adalah kemunculan dari fitur dan dari kategori.

Tabel 1. Tabel Kontingensi Seleksi Fitur Chi Square

	$e_c = 1$	$e_c = 0$
$e_t = 1$	N_{11}	N_{10}
$e_t = 0$	N_{00}	N_{11}

Langkah terakhir dalam tahapan seleksi yaitu mengurutkan semua hasil dari perhitungan seleksi fitur dari urutan yang terbesar sampai yang terkecil. Proses penghapusan fitur dilakukan apabila dalam penerimaan hipotesis independen terpenuhi.

3. Tahap klasifikasi data

Pada tahap klasifikasi data menggunakan Naïve Bayes dan kNN

4. Tahap evaluasi kinerja sistem

Dalam tahapan melakukan evaluasi sistem, perhitungan menggunakan tabel kontingensi [7] yang diberikan pada Tabel 2.

Tabel 2. Tabel Kontingensi Evaluasi Kinerja Sistem

#	Relevant	Not Relevant
Retrieved	True Positive (TP)	False Positive (FP)
Not Retrieved	False Negative (FN)	True Negative (TN)

III. HASIL DAN PEMBAHASAN

Data sebanyak 200 opini yang digunakan terdiri dari 100 opini positif dan 100 opini negatif. Sebagai data latih digunakan sebanyak 100 data yaitu data yang terbagi masing-masing 50 opini positif dan 50 opini negatif.

3.1. Hasil evaluasi kinerja sistem Algoritma Naïve Bayes

Hasil dari klasifikasi menggunakan algoritma Naïve Bayes mendapatkan nilai akurasi sebesar 75%, berikut gambaran tahap evaluasi sistem menggunakan table kontingensi.

Tabel 3. *Positive_predictive_value: unknown* (positive class: Tidak)

	true Ya	true Tidak	class precision
pred. Ya	3	1	75.00%
pred. Tidak	0	0	0.00%
class recall	100.00%	0.00%	

Tabel 4. *Negative_predictive_value: 75.00%* (positive class: Tidak)

	true Ya	true Tidak	class precision
pred. Ya	3	1	75.00%
pred. Tidak	0	0	0.00%
class recall	100.00%	0.00%	

Tabel 5. *Precision: unknown* (positive class: Tidak)

	true Ya	true Tidak	class precision
pred. Ya	3	1	75.00%
pred. Tidak	0	0	0.00%
class recall	100.00%	0.00%	

Tabel 6. *Recall: 0.00%* (positive class: Tidak)

	true Ya	true Tidak	class precision
pred. Ya	3	1	75.00%
pred. Tidak	0	0	0.00%
class recall	100.00%	0.00%	

3.2. Hasil evaluasi kinerja sistem Algoritma kNN

Tabel 7. *Precision* (positive class: Tidak)

	true Ya	true Tidak	class precision
pred. Ya	2	1	66.67%
pred. Tidak	1	0	0.00%
class recall	66.67%	0.00%	

Tabel 8. *Recall:* (positive class: Tidak)

	true Ya	true Tidak	class precision
pred. Ya	2	1	66.67%
pred. Tidak	1	0	0.00%
class recall	66.67%	0.00%	

Table 9. *Positive_predictive_value: 0.00%* (positive class: Tidak)

	true Ya	true Tidak	class precision
pred. Ya	2	1	66.67%
pred. Tidak	1	0	0.00%
class recall	66.67%	0.00%	

Tabel 10. *Negative_predictive_value: 66.67%* (positive class: Tidak)

	true Ya	true Tidak	class precision
pred. Ya	2	1	66.67%
pred. Tidak	1	0	0.00%
class recall	66.67%	0.00%	

IV. KESIMPULAN

Analisis pengguna Whasapp yang dilakukan dalam proses klasifikasi setelah melewati tahapan pengujian dan pelatihan data, serta evaluasi maka dapat disimpulkan pada penggunaan Metode Naive Bayes dapat dipilih sebagai metode klasifikasi pada karakteristik data dikarenakan prosesnya yang sederhana dan akurat, dapat mengoptimalkan hasil dari klasifikasi. Dengan Jumlah data sebanyak 200 data nilai rata-rata akurasi yang didapat sebesar 75%. Di bandingkan menggunakan metode klasifikasi lainnya yaitu kNN yang

memiliki akurasi sebesar 66,7%, maka metode Naïve Bayes dapat digunakan pada karakteristik data serupa.

Adapun faktor yang menjadi penghambat dan mempengaruhi hasil dari akurasi, yaitu data latih yang kurang kompleks karena jumlah pengguna didalam grup tersebut tidak terlalu banyak pengguna aktif dan penyebaran dari jenis kata yang tidak merata dengan kata lain serta jumlah dari jenis kata tertentu pada data latih dokumen yang positif digunakan tidak sama dan kurang seimbang dikarenakan ada dominasi percakapan yang dilakukan dari beberapa pengguna saja, sebaiknya percakapan lebih beragam dan tidak didominasi salah satu pengguna tertentu.

DAFTAR PUSTAKA

- [1] K. Nistanto, "Negara dengan Jumlah Pengguna WhatsApp Terbanyak di Dunia," *Kompas*, 17 November 2016.
- [2] Hidayat, Andi Nurul. "Analisis Sentimen Terhadap Wacana Politik Pada Media Masa Online Menggunakan Algoritma Support Vector Machine Dan Naive Bayes." *Jurnal Elektronik Sistem Informasi dan Komputer* 1.1 (2015): 12-18.
- [3] Maning, C., Raghavan, P. & Schutze, H., 2008. *Introduction to Information Retrieval*. London: Cambridge University Press
- [4] Zhang, L., Ghosh, R., Dekhil, M., Hsu, M. & Liu, B., 2011. *Combining Lexiconbased and Learning-based Methods for Twitter Sentiment Analysis*, Chicago: Hewlett-Packard Development Company, L.P
- [5] Witten, I. H., Frank, E., & Hall, M. A. 2011. *Data Mining: Practical Machine Learning and Tools*. Burlington : Morgan Kaufmann Publisher
- [6] Vercellis, Carlo. 2009. *Business Intelligent: Data Mining and Optimization for Decision Making*. Southern Gate, Chichester, West Sussex : John Willey & Sons, Ltd
- [7] Virtanen, Seppo, and Mark Girolami. "Precision-Recall Balanced Topic Modelling." *Advances in Neural Information Processing Systems*. 2019.

